Precision Agriculture, 5, 435–444, 2004

Frequency Analysis of Yield for Delineating Yield Response Zones*

K. DIKER Kenan.diker@ars.usda.gov
D. F. HEERMANN AND
M. K. BRODAHL
USDA-ARS Water Management Unit, 2150 Centre Avenue, Building D, Suite 320, Fort Collins, CO 80526

Abstract. The yield in any given field or management zone is a product of interaction between many soil properties and production inputs. Therefore, multi-year yield maps may give better insight into determining potential management zones. This research was conducted to develop a methodology to delineate yield response zones by using two-state frequency analysis conducted on yield maps for 3 years on two commercial corn fields near Wiggins, Colorado. A zone was identified by the number of years that yield was equal and greater than the average yield in a given year. Classes producing statistically similar yield were combined resulting in three potential yield zones. Results indicated that the variability of yield over time and space could successfully be assessed at the same time without the drawbacks of averaging data from different years. Frequency analysis of multi-year yield data could be an effective way to establish yield response zones. Seventeen percent of the field #1 consistently produced lower yield than the mean while 43% of the field produced yield over the mean. Corresponding values for field #2 were 6% and 42%. The remainder of the fields produced fluctuating yields between years. These spatially and temporally sound yield response maps could be used to identify the yield-limiting factors in zones where yield is either low or fluctuating. Yield response maps could also be helpful to delineate potential management zones with the help of resource zones such as electrical conductivity and soil maps, along with the directed soil sampling results.

Keywords: two-state frequency analysis, management zones, yield response zones, corn

Introduction

Variable rate application of agricultural inputs to crop requires well-defined nutrient management zones—a part of a field with similar nutrient supplying capabilities. Moore and Wolcott (2000) stated that management zones for production fields might lead to increased efficiency in application of inputs based on yield.

Management zones can be created by either qualitative data such as bare soil color or quantitative data such as soil electrical conductivity or soil organic matter. Employing these soil variables for delineating management zones by themselves may lead to overlooking the interactions between these variables and the environment that drives the final yield. As stressed by Stafford *et al.* (1999), gathering information on the interactions is possible, but is time and labor intensive.

*Mention of trademark or vendor does not constitute a guarantee or warranty of the product by the USDA and does not imply any exclusion of other products or vendors that may also be suitable.

Yield mapping is a simple and inexpensive way to monitor the integrated effect of the interactions of factors affecting yield. Nolan *et al.* (1996) stated that a yield map was an important tool in decision-making to manage the soil fertility at the sub field scale.

Yield maps will not always translate directly into nutrient management zones, because yield response is a "bio-assay" of crop environments in the field. The crop responds to the integrated environment and, as a living organism, the plant has some ability to compensate for different environmental conditions. A crop's response to a particular environmental factor can be non-linear, related to threshold levels, and/or influenced by the state of other factors. Since the plant responds to an integrated environment, similar crop responses could be observed over different crop environments and the management required to sustain or improve a crop may differ even though the yield patterns are the same. In some fields, there may be a single dominant driving factor for yield, such as salinity, while in other cases, the crop responds to a suite of interacting factors. The dominance of a factor or set of factors may change with different growing season conditions and/or management practices, and even within different areas of a field in the same year. Site-specific management based solely on yield zones may not successfully target the factors driving yield. However, yield zones could be important tools for the development of site-specific management strategies. Yield zones can help define yield goals for an area of the field. Coupled with producer knowledge, yield zones may be a good starting point for developing site-specific management strategies. The addition of other related layers of site-specific information could strengthen the overall management strategy. Moore and Wolcott (2000) stated that yield zones could be transformed into nutrient management zones through their use in defining yield goals.

One technique for identifying yield based response zones is averaging relative yield data of multiple years. However, averaging data would neutralize some differences occurring over years, leading to loss of valuable temporal information. Moore and Wolcott (2000) used an averaging relative yield technique that overcame the problem of crop-to-crop differences. They presented a yield map created by overlaying three-year normalized yield averages. Taylor and Whitney (2001) took the same approach, but averaged 2 years of normalized yield data to assess the possibility of using yield monitor data to direct soil sampling in comparison with grid sampling. They reported that yield-monitor-directed sampling was certainly an option as a lower cost alternative to grid sampling. However, it was unlikely to describe as much spatial variability as grid sampling. This could be due to averaging the two-year standardized yield that would have smoothed out some of the temporal changes. Lark and Stafford (1996) developed a methodology to interpret spatial and temporal variability of multi-year yield data by pattern recognition. Patterns identified by fuzzy cluster analysis indicated that temporal variation patterns could be established. Panneton *et al.* (2001) successfully used a fuzzy classification to delineate stable high, average and low yield zones. Blackmore (2000) developed a spatial trend map by averaging the yield in each grid cell over the years. Multi-year yield data were effectively categorized using the spatial trend map in conjunction with the coefficient of variation. Blackmore *et al.* (2003) used the yearly mean to determine if the individual yield values were above or below the mean. This was used with stability in time to preset the yield classes in developing spatial trend maps. The technique used an arbitrary 1 t/ha stability threshold for determining temporal stability. When the threshold was reduced from 1 to 0.8 t/ha, the unstable areas increased from 7% to about 50% of the field. This indicates that the threshold played a significant role in the yield zone classification in this technique. The use of a threshold could be avoided by introducing two-state frequency analysis for generating aggregated multi-year yield maps into one map. Such an analysis could also provide producers with information on the areas where similar yield levels were achieved repeatedly over the years. This might help delineate stable low or high yielding areas that could be useful in decision making in future years. Frequency analysis naturally requires standardization of the yield such as "high" and "low" yield at any location within the field. Such normalization compensates for both temporal and crop-to-crop differences (Diker *et al.*, 2002).

This study was undertaken to develop a methodology to delineate yield response zones by applying two-state frequency analysis to multi-year yield maps and to consider the use of these yield zones in developing site specific management strategies.

Materials and methods

Materials

The yield data used in this study were from a precision agriculture study conducted on two center pivot irrigated corn fields in northeastern Colorado, USA. Data collection was carried out in years 1997–2000. However, 1998 field #1 and 2000 field #2 data were not included in the analysis because of heavy hail damage.

Field #1 was 70.8 ha; soils included Valentine sand (Sandy, mixed nonacid, mesic Typic Ustipsamment), Bijou loamy sand (Course loamy, mixed, mesic Mollic Haplargid), and Truckton loamy sand (Course loamy, mixed, mesic Udic Argiustoll) (Soil Survey Staff, 1996); field #2 was 52.6 ha with similar soil types.

The corn yield was harvested by combines equipped with yield monitors and a Global Positioning System (GPS) with differential correction. Yield data were processed and mapped with FarmHMS (Red Hen Systems Inc.) using MapInfo software. Yield information were imported into the geographic information system (GIS) package ArcView 3.2 for further analysis.

Methods

Yield monitor data were cleaned by removing yield amounts less than 2.2 t/ha and greater than 18.8 t/ha. These numbers were selected to eliminate/minimize the skewness in the yield after careful evaluation of yield distribution. Yield higher than 18.8 t/ha seemed unrealistic and was probably related to grain flow lag when the combine was stationary. Yield lower than 2.2 t/ha are from the areas near the end of the combine pass where grain is being harvested, but has not yet reached the monitor. This process removed apparent low yields occurring at the edge of both fields and removed some of the skewness caused by those low yield amounts in field

#1 data set. However, it did not remove the skewness in yield data of field #2. The mean of the yield in each year for each field was calculated from these cleaned data. The yield data were interpolated by using inverse distance weighting (IDW) in ArcView 3.2 (ESRI, 1996) with 12 nearest neighbors and a grid cell size of 6.1 m that corresponded to harvester width. The resulting grid maps were reclassified by assigning the state 0 to yield below the within-year-mean and 1 to yields at and above the mean. Consequently, two standard yield states were created to use in the frequency analysis for all years and fields.

An ArcView script of two-state frequency analysis developed by Zimmerman (2000) was imposed on the yield data from 3 years for both fields. This analysis checks every grid cell and determines how many years that cell was in state 1. The number of the years that yield was above-mean (i.e. state 1) was assigned to that 6.1 m grid cell. The resulting two-state frequency maps showed the number of years of at and above mean yield in each grid cell. These maps were too detailed for practical use in managing crop inputs, therefore, a smoothed frequency map of the yield was created by filtering the maps using ArcView's Neighborhood function. This function uses a moving window of 9×9 cells which passes over the centers of all grid cells, and assigns the "majority" value of the 81 cells to the center grid cell.

The zones that statistically had similar yield means as indicated by box and whiskers plots and the zone means over the years were combined. The yield frequency zones with overlapping boxes in box and whiskers plots were considered similar and combined.

Results and discussion

Figure 1 shows the reclassified yield maps for both fields over 3 years. Maps indicated that spatial distribution of crop yield changed over the years. The percentages of below-mean yield areas decreased from 42% in 1997 to 37% and 35% in 1999 and 2000, respectively in field #1 (Figure 1). This decrease could have resulted from some fertilizer experiments leading to higher yield within the experimental sites. The fertilizer experiment areas that would interfere with the yield results were kept in the analysis due to the fact that these areas were small compared to the whole field. In addition, any technique dealing with yield data that is affected by many factors should be robust to some unusual conditions. In field #2, the below-mean yielding areas were 31%, 43% and 37% for 1997, 1998 and 1999, respectively.

Below-mean yielding areas were located in the north-east and south-west of field #1 and mainly in the south-east half of field #2. Visual comparisons of the maps in Figure 1 indicated that the distribution pattern of the above and below mean yielding areas was generally consistent spatially with the exception of field #2 in 1997 which was a wet year. This year and site had 31% of area below-mean dispersed across the field as compared to the more spatially aggregated areas of below-mean yield in 1999 and 2000. The yield consistency might be an important factor when translating yield zones into fertilizer management zones.

Frequency analysis was run on the maps in Figure 1 to quantify the consistency of the above and below-mean yielding areas. The resulting frequency maps for each

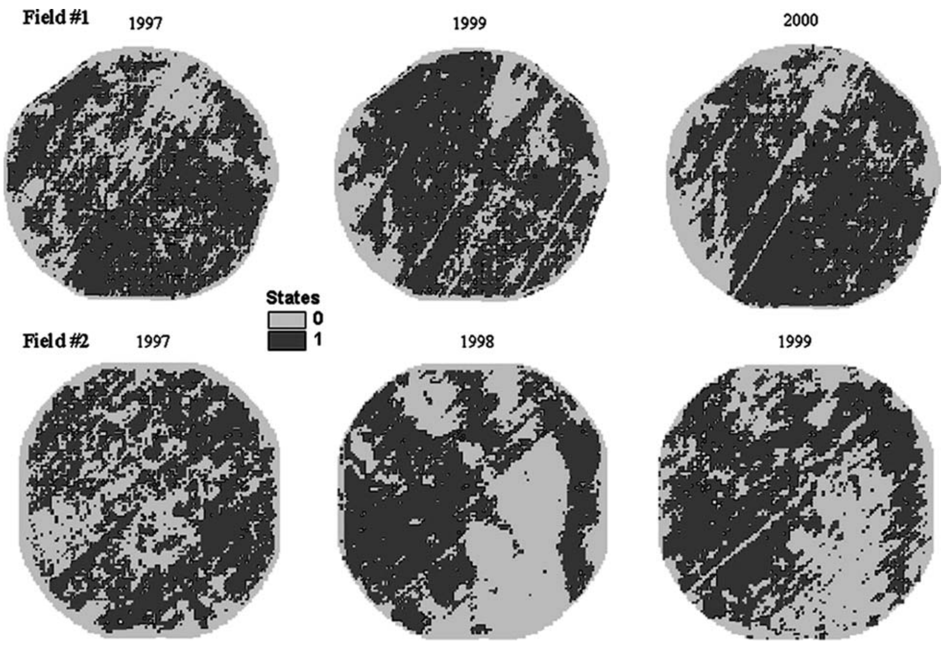

Figure 1. Reclassified yield maps for both fields showing distribution of below (0) and above (1) within-year mean states.

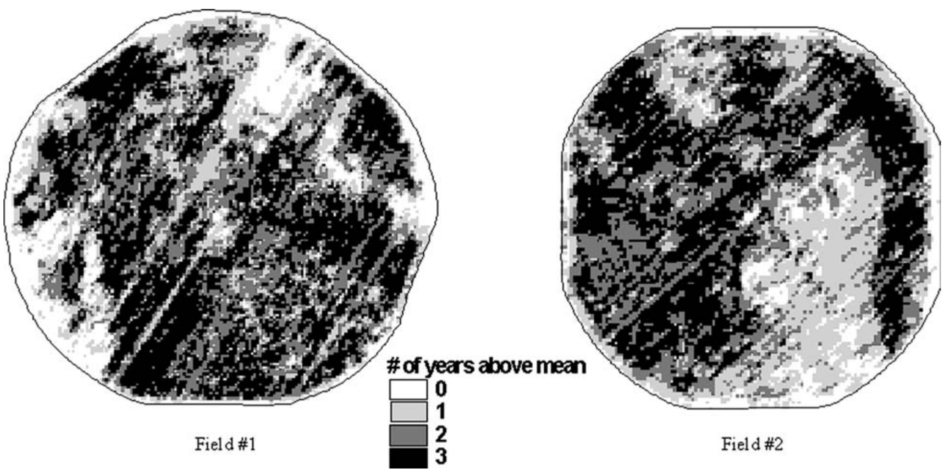

Figure 2. Fields #1 and #2 maps showing distribution of number of years out of 3 for the above-mean yield state.

field are presented in Figure 2. Four yield response zones were developed. The yield frequency class 0 shows areas where state 1 (above-mean yield) never occurred meaning these areas produced consistently below-average yield. The yield frequency

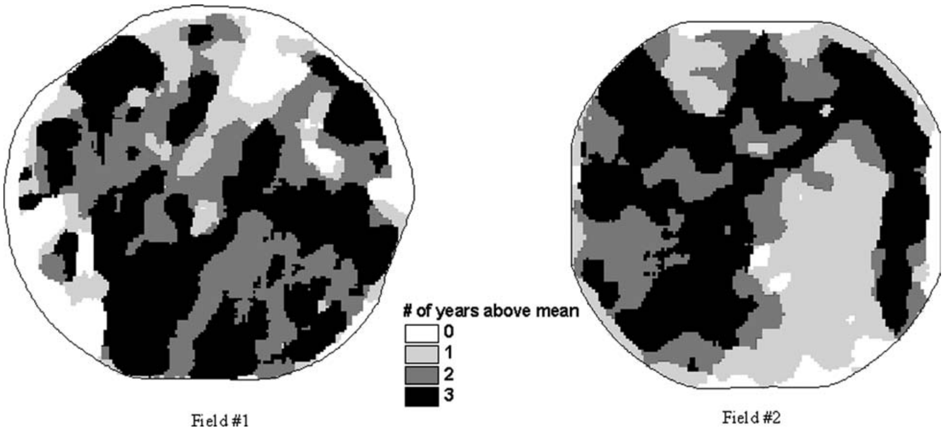

Figure 3. Generalized maps of maps shown in Figure 2.

classes 1 and 2 represent areas where state 1 was encountered 1 and 2 years out of 3 above-mean yield state. The areas represented by yield frequency classes 3 had above-mean yield 3 out of 3 years of study suggesting that these areas are consistent in producing above-mean yield. The yield frequency classes labeled 1 and 2 show the fluctuating yield response that would require a better management to produce consistent yield over the years, assuming the yield response is related to manageable factors. In field #2, however, below-mean yielding areas localized around the perimeter of the field and the south-east half of the field where the soil was sandy.

After aggregating the 6.1 m $\times$ 6.1 m grids by using ArcView's neighborhood function, filtered yield frequency zones were obtained and are shown in Figure 3. Frequency of above-mean map of field #1 in Figure 3 demonstrated that 17% of the field consistently produced lower yield than the mean (yield frequency class 0) while 43% of the field produced yield over the mean (yield frequency class 3) every year. Yield in the rest of the field fluctuated from year to year. Corresponding percentages for field #2 were 6% and 42%. A relatively stable 0-1 pattern of variability from year to year was observed as opposed to the findings of Stafford *et al.* (1999) where yield patterns were not stable from year to year possibly due to the different patterns and range of yield variability of rain-fed crops from year to year. Mean yields of each frequency class and box and whisker analysis were used to combine statistically similar frequency classes. Mean yield values in each frequency class given in Table 1 showed that mean yield in zones 1–3 were similar in 1997 and 1999 whereas, in 2000, classes 2 and 3 produced higher yield than other classes in field #1. This result meant that yield frequency classes 2 and 3 in field #1 should be combined. Box plots given in Figure 4 showed that yield frequency classes 1 and 2 and 2 and 3 in field #1 were overlapping. However there was no overlap between frequency classes 1 and 3. These results indicated that there was no clear separation between the yield frequency classes 1 and 2 and 2 and 3. Furthermore, the spatial orientation of zone 1 indicated that it was a transitional zone between zone 0 and zones 2 and 3. The largest zone 1 was located between zones 0 and 2 in the north of the field. Therefore, frequency classes 1 and 2 could be combined for the sake of yield stability. The yield stability

Table 1. Mean yield in each zone for fields 1 and 2. Zones are based on how many years out of three were above the within year average yield

Zone	Average Yield (kg/ha)		
	1997	1999	2000
Field #1			
0	8.9	8.6	9.6
1	10.1	11.0	11.1
2	10.6	11.4	12.3
3	11.2	11.8	12.7
	1997	1998	1999
Field #2			
0	9.5	9.2	12.0
1	12.0	11.1	13.3
2	12.2	13.2	14.7
3	12.8	14.0	15.1

measures how well the yield zones maintain their relative yield ranking from season to season. The stability must be high for yield zones to be useful in management (Moore and Wolcott, 2000). By combining yield frequency classes 1 and 2, three zones for field #1 were identified; one of stable high yielding (3 years of above-mean), one of stable low yielding (0 year of above-mean) and unstable yielding zones (1 and 2 years of above-mean). As for field #2, the last 2 years of means (Table 1) and box analysis suggested that yield frequency classes 2 and 3 in field #2 could be combined.

Figure 5 presents the final yield zones showing yield variability over time and space created by yield mapping. As seen in Figure 5, 43% and 66% of the fields #1 and 2, respectively, were classified as high yield zone. The above-mean zones in field #1 had a more disconnected, patchy distribution whereas in field #2, there was a single unit. Medium and low frequency zones were patchy in both of the fields. The low yield frequency areas located at the perimeter of both fields indicated that special attention should be given to perimeter areas under center pivot irrigated fields. Low yields at the perimeter could partially be due to yield monitor errors at start and stop of the run. The measurements of the length of the low yielding areas at the edges ranged from about 25 to 85 m in field #1 and 25 to 71 m in field #2. The length seemed to be longer than experimental artifact. Therefore, it was concluded that low yielding areas at the perimeter resulted mostly from the relatively higher sand content of the soil at the perimeter.

The frequency analysis of multi-year yield data produced spatially and temporally sound yield response maps. It was especially effective to monitor yield variability over the field between the years since averaging between years that smooths out the year to year variability was not involved. Using these yield response maps, the producers and/or consultants with a few years of yield data could identify the yield-limiting factors in each zone then the problem could be corrected in subsequent years. Based on the yield zone information, management changes could be

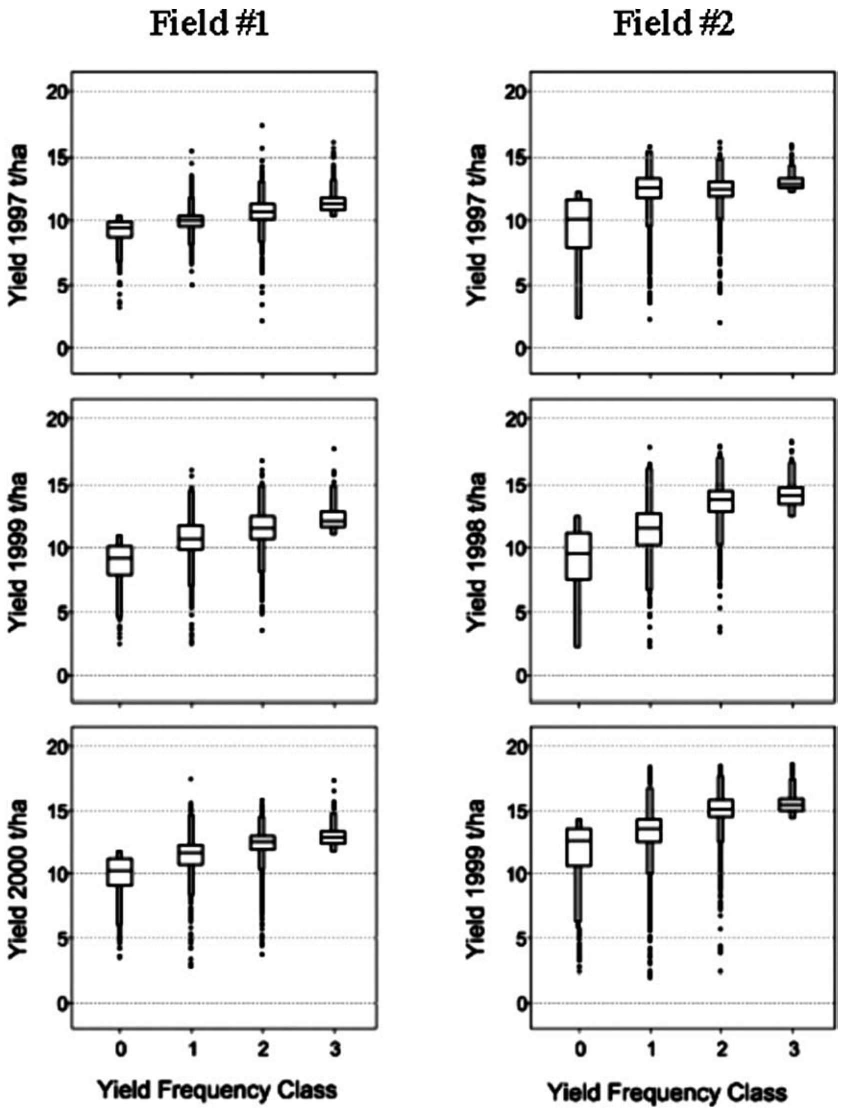

Figure 4. Box and Whisker plots of yield, by year, within each yield frequency class. Lower boundary of box is at the first quartile (25%), upper boundary at the third quartile (75%), the horizontal bar at the median (50%). The ''whiskers'' are, at a maximum, 1.5 times the length of the box and outliers (data points beyond the whiskers) are shown as points.

made to produce an optimum yield in the areas where the yield was fluctuating because the center pivot irrigated agriculture practiced in eastern Colorado allows easy management changes due to fertigation and chemigation capabilities. The maps could also be helpful to the users to delineate potential management zones with the assistance of tools such as soil electrical conductivity and soil survey maps.

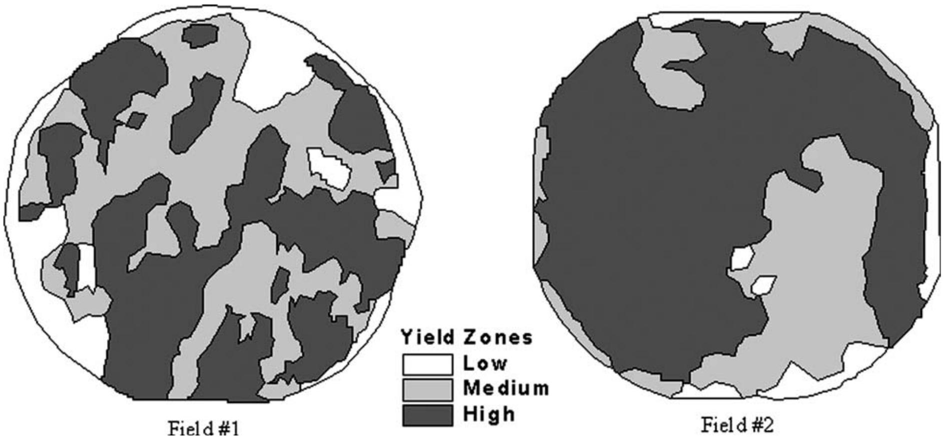

Figure 5. Yield zones for fields #1 and #2.

Conclusions

This study demonstrated that frequency analysis of multi-year yield data is a promising way to delineate yield response zones in a field. The variability of yield over time and space could be assessed without the drawbacks of averaging data from different years. The technique differs from previously reported techniques since this technique does not require preset final yield classes and uses statistically sound, two-state frequency analysis. Results indicated that frequency analysis of multi-year yield data could be an effective way to establish yield response zones. Seventeen percent of field #1 consistently produced lower yield than the mean while 43% of the field produced yield over the mean. Corresponding values for field #2 were 6% and 42%. The remaining areas within the fields produced fluctuating yield from year to year. Box and whiskers plots seemed to be useful to combine statistically similar areas with similar means and data distribution. These spatially and temporally sound yield response maps might be useful in identifying zones in which investigation into manageable factors for yield improvement could be implemented. They could also be helpful to delineate potential management zones with the help of tools such as electrical conductivity and soil survey maps.

References

Blackmore, S. 2000. The interpretation of trends from multiple yield maps. Computers and Electronics in Agriculture **26**(1), 37–51.

Blackmore, S., Godwin, R. J. and Fountas, S. 2003. The analysis of spatial and temporal trends in yield map data over six years. Biosystems Engineering **84**(4), 455–466.

Diker, K., Buchleiter, G. W., Farahani, H. J., Heermann, D. F. and Brodahl, M. K. 2002. Frequency analysis of yield for delineating management zones. In: *CD Proceedings of the* 6th *International Conference on Precision Agriculture and Other Precision Resources Management,* edited by P. C. Robert (ASA-CSSA-SSSA, Madison, WI, USA).

ESRI, 1996. *Working with the ArcView spatial analyst* (Environmental System Research Institute, Redlands, California).

Lark, R. M. and Stafford, J. V. 1996. Consistency and change in spatial variability of corn yield over successive seasons: methods of data analysis. In: *Proceedings of the* 3rd *International Conference on Precision Agriculture*, edited by P. C. Robert, R. H. Rust and W. E. Larson (ASA-CSSA-SSSA, Madison, WI, USA), pp. 141–149.

Moore, S. H. and Wolcott, M. C. 2000. Using yield maps to create management zones in field crops. Louisiana Agriculture **43**(3), 12–13.

Nolan, S. C., Haverland, G. W., Goddard, T. W., Green, M., Penney, D.C., Henrikson, J. A. and Lachapelle, G. 1996. Building a yield map from geo-referenced harvest measurements. In: *Proceedings of the* 3rd *International Conference on Precision Agriculture*, edited by P. C. Robert, R. H. Rust and W. E. Larson (ASA-CSSA-SSSA, Madison, WI, USA), pp. 885–892.

Panneton, B., Brouillard, M. and Pikutowski, T. 2001. Integration of yield data from several years into a single map. In: *Proceedings of the* 3rd *European Conference on Precision Agriculture*, edited by G. Grenier and S. Blackmore (Agro Montpellier, Montpellier, France) pp. 73–78.

Soil Survey Staff, 1996. *National Soil Survey Handbook*. *Title 430-VI*. (USDA Natural Resources Conservation Service. U S Government Printing Office, Washington, DC. 20250).

Stafford, J. V., Lark, R. M. and Bolam, H. C. 1999. Using yield maps to regionalize fields into potential management units. In: *Proceedings of the* 4th *International Conference on Precision Agriculture*, edited by P. C. Robert, R. H. Rust and W. E. Larson (ASA-CSSA-SSSA, Madison, WI, USA), pp. 225–237.

Taylor, R. and Whitney, D. 2001. Using yield monitor data to direct soil sampling. http://www.oznet.ksu.edu/pr_prcag/yield-moniter.shtml. Accessed on March 13, 2001.

Zimmerman, K. 2000. Two-State Markov chain. http//arcscripts.esri.com/details.asp?dbid = 11483. Accessed on January 20, 2002.